16.7 Comparison of Split- Versus Connected-Core Supplies in the POWER6™ Microprocessor

[1]Norman James, [2]Phillip Restle, [1]Joshua Friedrich, [3]Bill Huott, [1]Bradley McCredie

[1]IBM, Austin, TX
[2]IBM T.J. Watson, Yorktown Heights, NY
[3]IBM, Poughkeepsie, NY

POWER6™ is a dual-core microprocessor fabricated in a 65nm SOI process with 10 levels of low-dielectric copper interconnects. The die, shown in Fig. 16.7.1, measures 341mm^2, contains over 700M transistors, delivers clock frequencies exceeding 5GHz in high-performance applications, and consumes less than 100W in power-sensitive applications [1]. Chips with split and connected core power supplies are fabricated, modeled, and tested, showing both the advantages and disadvantages of each, with important implications for chips with large numbers of cores. One of the power grid designs has the two processor cores on isolated logic power boundaries. The other design has both cores tied into the rest of the chip (called the nest) on both the chip and package.

There are advantages and disadvantages for each of the two power grid designs. The split cores allow for independent voltage-tuning optimizing power versus performance. The manufactured die has systematic and non-systematic variation across the chip that can make one core faster and have higher leakage than the other, but both cores run at the same clock frequency on POWER6. With separate power domains, the voltage can be lowered on the core with faster circuits. This is done on previous generations of PowerPC microprocessors [2]. Another advantage of split cores is supporting power down modes for an unused core. The disadvantage of the split power grid is that the cores do not benefit from being connected with the relatively quiet nest. The cores consume considerably more power and have much higher dI/dt than the nest, which is made up of mostly level-2 cache and I/O. With cores and nest connected, the cores get the benefit of sharing the quiet on-chip nest capacitance and also share a lower-inductance path to the package decoupling capacitors, further reducing power noise in the cores.

Figure 16.7.2 is a simple schematic illustrating the mid-frequency characteristics of the chip and package power distribution. In the figure, R1, C1, L1, and R2 represent the package, which, for the POWER6 package, has a 125MHz resonant frequency. C2 is the intrinsic wire and device capacitance. R3 and C3 represent the added on-chip decoupling capacitance. R4 represents the (nonlinear) leakage and R5 is used to cause the current step in simulation. A single POWER6 core is capable of causing a 13W power step within about 20 clock cycles. Detailed chip-package simulation predicts this causes a 130mV power droop when the cores are split. When the cores are tied, this droop is cut in half.

Directly measuring the on-chip noise is difficult with external measurement equipment because there is limited visibility to the internal power grid. POWER6 has only one external voltage sense per core and the bandwidth is limited to around 2GHz. To supplement this, two on-chip measurement circuits are used: skitter [3] and critical path monitor [4] (CPM) circuits. Both circuits use the similar principle of a delay-line with latches tapped off after each delay step. Both circuits can be used in a *sticky* mode to find the worst case, or in a *single sample* mode (or oscilloscope mode) such that the single-cycle delay is measured for any desired cycle.

Figure 16.7.3 shows the number of latches changing state (latch activity) and delay per inverter measured by the skitter circuit for 1000 cycles at 4GHz for a split-core chip. Increased latch activity causes power supply droops, resulting in an increase in the delay per stage measured by the skitter circuit. The graph shows that the switching activity increases from 1000 latches switching to 7000 latches in 18 core cycles. This is an increase in activity from 2% to 14% of the 50,000 total core latches. Power calculations predict this corresponds to a 13W power increase, which is used in the noise simulations. The corresponding skitter delay per stage, which closely follows the switching activity, shows a 17% change in delay. The simulated results using the model described above are also shown in Fig. 16.7.3 and have good agreement with the hardware.

Since both skitter and CPM delays are dependent on both V_{DD} and the clock period, we must eliminate the possibility that the results in Fig. 16.7.3 are caused by short cycles rather than V_{DD} droops. Fortunately on POWER6, we have the ability to send the core clock to a skitter in the nest and vice versa. Figure 16.7.4 shows that when this is done, the noise event is not seen in the nest skitter and is still seen in the core skitter, so the conclusion is that the 17% change is predominantly due to power supply noise. A 17% change in delay equates to a 200mV droop at 1.1V, which is consistent with simulation. The low jitter even in the presence of supply noise is achieved with the optimized low-latency transmission-line design of the clock distribution, whose latency is measured to be 4.9× less sensitive to V_{DD} variations than an inverter chain [5].

Next, we compare split- versus connected-core designs. Figure 16.7.5 compares the cycle-by-cycle trace of skitter delay per stage for the two power grid designs for several parts experiencing a 13W power change in one core. The tied power grid shows roughly half the noise of the split power grid, which is also consistent with the simulation. The connected power grid is the best choice for this case. However if both cores are shorted to the nest, a new concern arises that the noise from one core could travel to the other. In the worst case, the noise from the one core might in rare cases arrive just when the second core is experiencing a locally produced V_{DD} droop, resulting in a core-core interaction we call a perfect storm. Figure 16.7.6 shows the noise from one core traveling across the chip to the second quiet core, arriving 4ns later with approximately 1/3rd amplitude. For the POWER6, where the nest is relatively quiet and only two cores exist, which are spread apart, the noise is significantly attenuated by the time it reaches the other core, so that the core-core interaction effect is small. For other chips, especially those with more than two cores, shorting the power grids may not be the best option. For example on a four-core chip, a rare event could occur where 3 cores would send noise droops arriving at a 4th core at just the wrong time, producing a perfect storm.

While the results above on specific test patterns strongly imply the connected power design is best for this chip, it must also be proven that the connected core improves the maximum operating frequency (f_{max}) across all workloads. Because not all chips are the same, a large sample is required and the performance sort ring oscillators (PSRO) versus f_{max} curves are compared. Figure 16.7.7 shows that the connected power chips exhibit a 3-5% f_{max} improvement consistent with modeling and the detailed noise measurements.

References:

[1] J. Friedrich, B. McCredie, N. Jame, et al., "Design of the POWER6™ Microprocessor," *ISSCC Dig. Tech. Papers,* pp. 96-97, Feb., 2007.
[2] E.B. Cohen, N.J. Rohrer, P. Sandon, et al, "A 64B CPU Pair: Dual- and Single-Processor Chips," *ISSCC Dig. Tech. Papers*, pp. 106-107, Feb. 2006.
[3] P. Restle et al., "Timing Uncertainty Measurements on Power 5 Microprocessor," *ISSCC Dig. Tech. Papers*, pp. 354-355, Feb 2004
[4] A. Drake, R. Senger, H. Deogun, et al., "A Distributed Critical Path Timing Monitor for a 65nm Power Microprocessor," *ISSCC Dig. Tech. Papers,* pp. 398-399, Feb., 2007.
[5] M.G.R. Thomson, P.J. Restle, and N.K. James, "A 5 GHz Duty-Cycle Correcting Clock Distribution Network for the POWER6 Microprocessor," *ISSCC Dig. Tech. Papers*, pp. 384-355, Feb., 2006.

 1-4244-0852-0/07/$25.00 ©2007 IEEE.

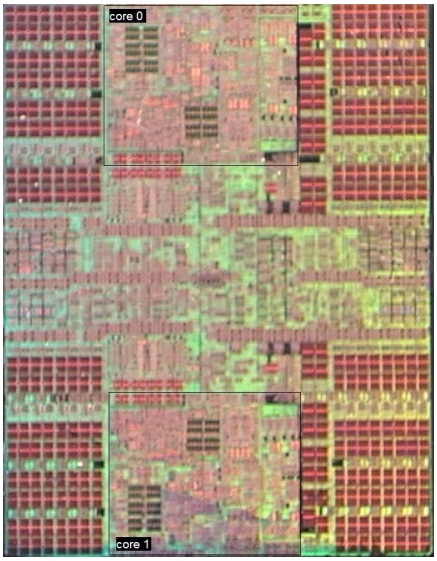

Figure 16.7.1: Die photo of POWER6 Microprocessor showing powerdomains.

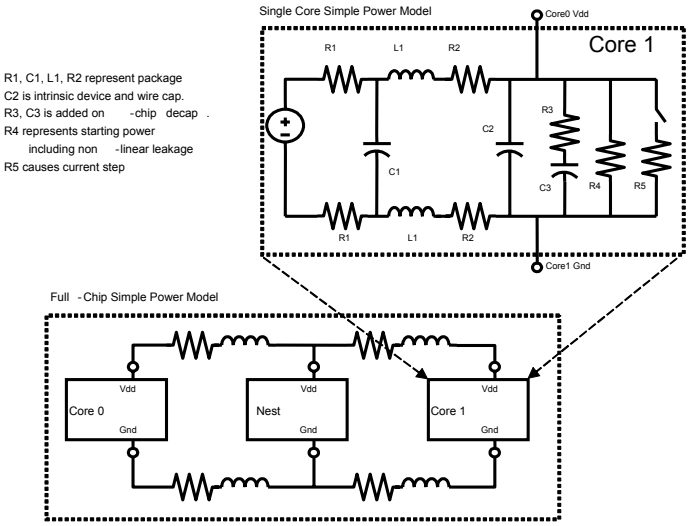

Figure 16.7.2: Simple schematic for a power distribution (split and shorted).

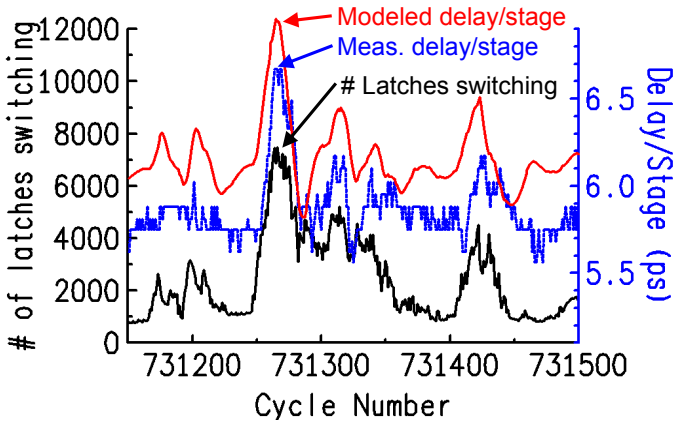

Figure 16.7.3: Cycle-by-cycle trace showing # of latches switching, skitter delay/stage,and the simulated delay/stage from power-supply modeling.

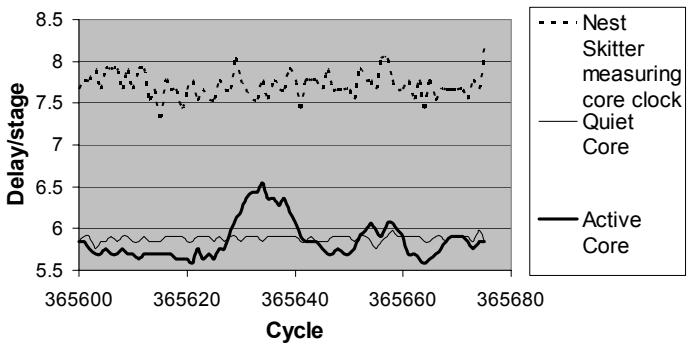

Figure 16.7.4: Cycle-by-cycle trace of skitter delay/stage showing clock is not a source of timing variation.

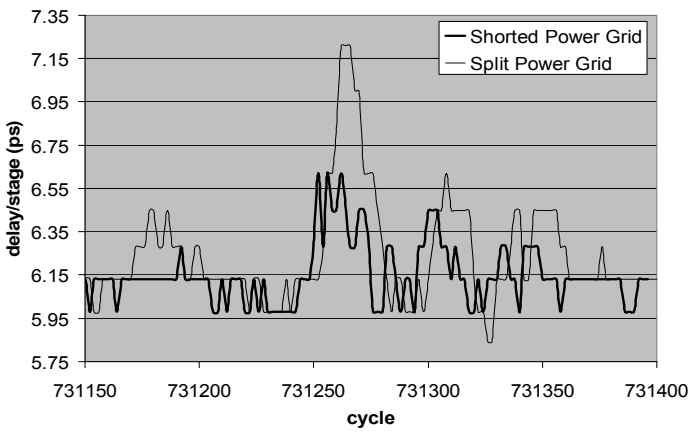

Figure 16.7.5: Cycle-by-cycle trace showing skitter delay/stage for split and shorted grids.

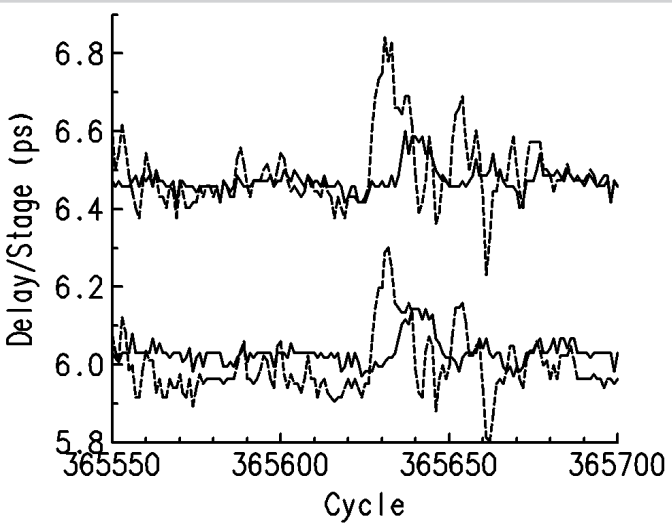

Figure 16.7.6: Cycle-by-cycle trace for two chips (A,B) with connected power supply showing noise traveling from source core to quiet core.

Continued on Page 604

16

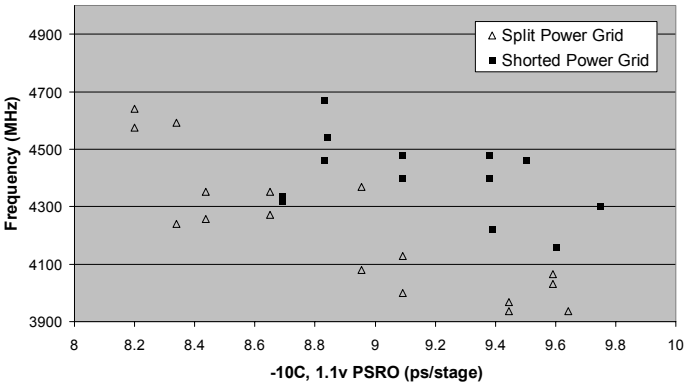

Figure 16.7.7, PSRO vs. functional fmaxfor split and shorted power grids (early POWER6 design).

 1-4244-0852-0/07/$25.00 ©2007 IEEE.